



Geneious Prime でシーケンス解析

第 26 回 リファレンスマッピングと SNP 検出 (その 1)

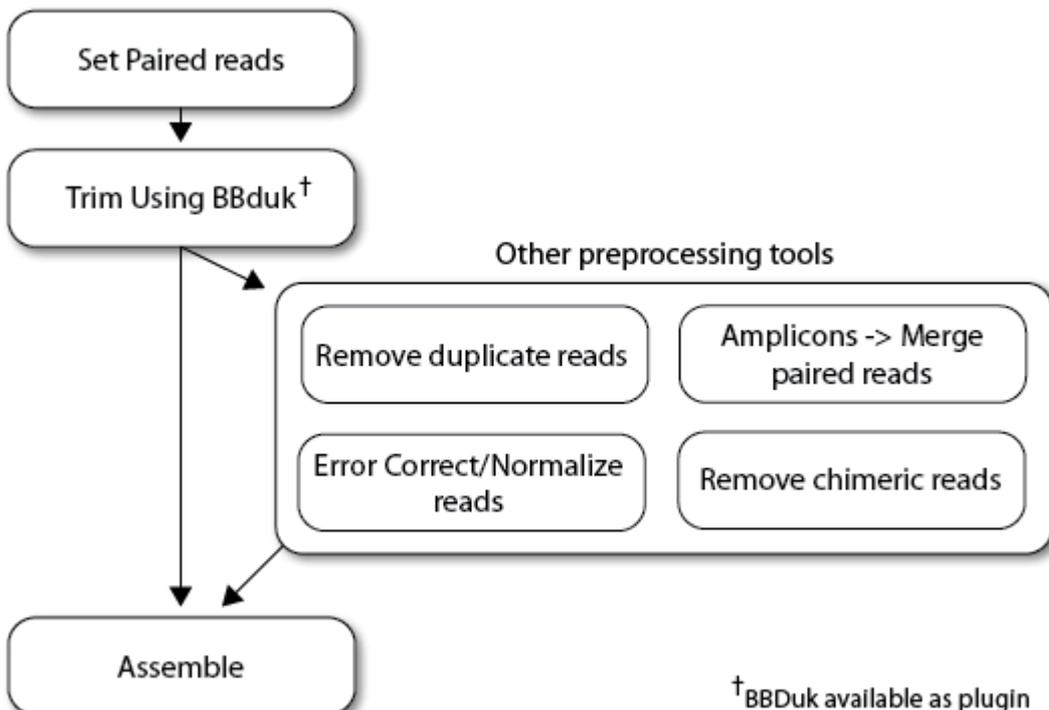


前回までの NGS リードの de novo アセンブリに引き続き、今回からは NGS リードをリファレンス配列にマッピングし、SNP を検出するための一般的なワークフローについてご紹介いたします。

今回もチュートリアル用データが用意されていますので、実際に操作して試してみたい方は、[こちらからダウンロード](#)し、ZIP ファイルを解凍せずにそのまま Geneious Prime にドラッグ & ドロップしてインポートしてください。

リファレンスマッピングも de novo アセンブリの場合と同じく、NGS リードの前処理を適切に行うことが非常に大切です。通常、マッピングでは de novo アセンブリに比べて必要な計算/時間コストが大幅に少ない場合が多いですが、マッピングの精度を向上させ、偽陽性 SNP の検出を減少させることに繋がります。

NGS リードの前処理ステップのフローは de novo アセンブリの場合と同様になりますので、[前回までの記事](#)をご参照ください。



チュートリアル用データの場合、*Escherichia coli* (大腸菌)ゲノム内の [yghJ 遺伝子](#)をリファレンス配列としてマッピングすることができる Illumina ショートリードのデータセット **yghJ paired Illumina reads** が用意されています。これはすでにペアリング済みのデータセットになっていますので、**Annotate & Predict** メニュー → **Trim using BBDuk** より、de novo アセンブリの場合と同様の設定でトリムを行ってください。

トリムが完了すると **yghJ paired Illumina reads (trimmed)** というファイルが新たに作成されます。

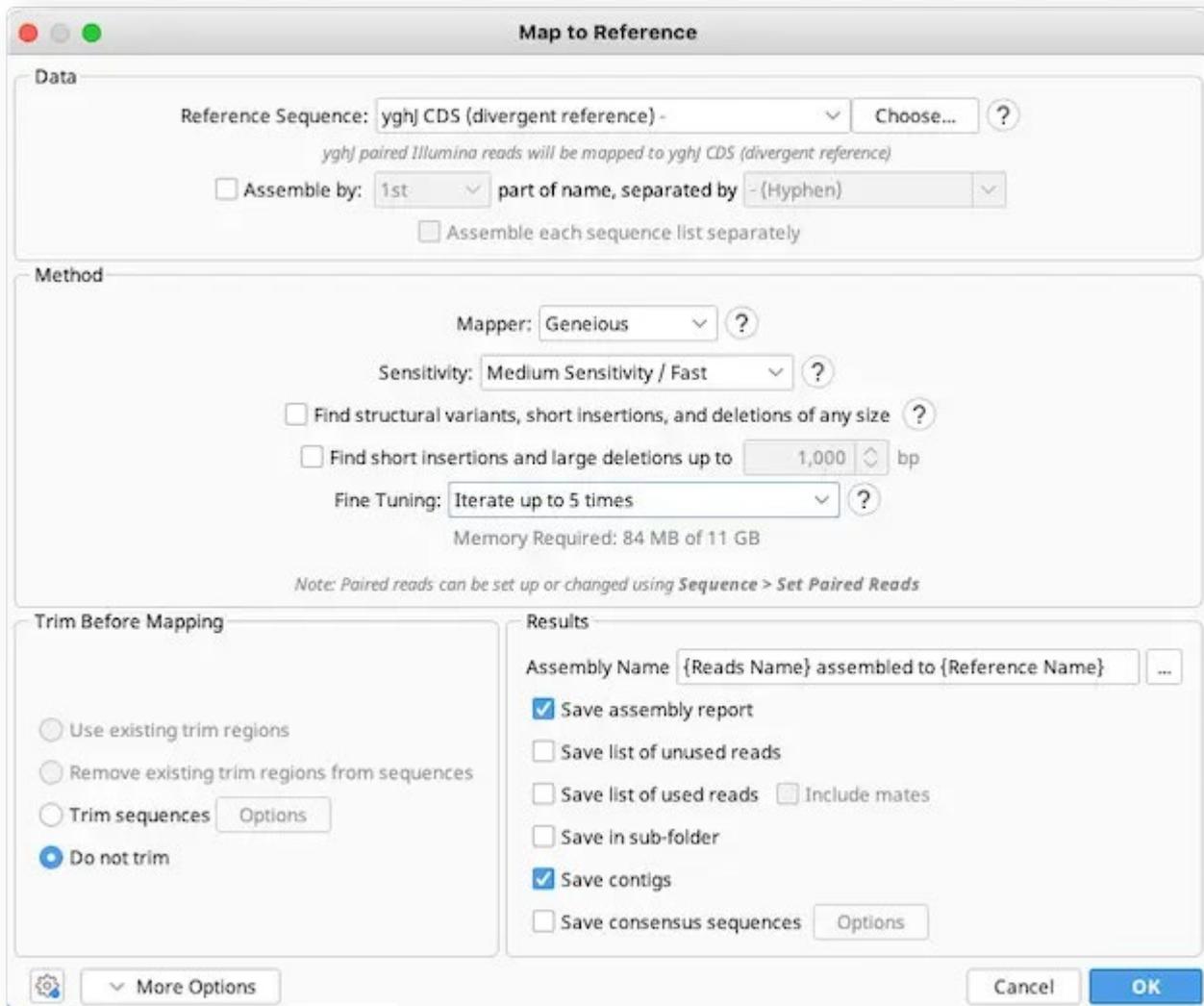
Shift キーを押したまま、トリム済みリードファイル **yghJ paired Illumina reads (trimmed)** とリファレンス配列 **yghJ CDS (divergent reference)** を選択し、**Align/Assemble** ボタン → **Map to Reference** コマンドをクリックします。

Geneious Prime は、選択された配列から、どの配列がリファレンス配列であるかを自動的に判別します。チュートリアルデータの場合、yghJ CDS (divergent reference) が Reference Sequence ボックスに表示されているはずですが、ドロップダウンメニューまたは **Choose** ボタンでデータベース内の任意の配列に変更することができます。

Method パネルで、マッピングアルゴリズムを選択することができます。チュートリアルの場合にはデフォルトの **Geneious** を選択します。マッピングアルゴリズムにはそれぞれ長所短所があり、**?** ボタンをクリックすることで詳細を確認することができますので、マッピングするデータの種類の適したものを選択することができます。

Sensitivity は、**Medium Sensitivity/Fast** に設定します。Geneious Prime はデータセットに応じて適切な感度を自動的に選択してくれます。十分なカバレッジがある NGS データの場合、High Sensitivity を使用しても時間がかかるだけでマッピングの結果は改善されにくいいため、Medium または Medium-Low Sensitivity をお勧めします。**Fine Tuning** オプションは、マッピング後にコンセンサス配列にマッピングする処理を繰り返すことでマッピング結果を向上させるものです。通常、デフォルト設定の **Iterate up to 5 times** のままで良いと思います。

Trim Before Mapping パネルでは **Do not trim** を選択します。NGS リードの場合、トリムはここで設定せず、事前に BBDuk で行います。**Results** パネルでは **Save contigs** を必ず選択します。**Save assembly report** はマッピングのレポートを確認した場合、**Save in subfolder** はファイルをサブフォルダに保存したい場合に選択します。



この設定で **OK** をクリックしてマッピングを実行すると、ドキュメントテーブルに、リファレンスにマッピングされたリードの **Contig** と、**Assembly Report** の 2 つの新しいドキュメントが作成されます。Assembly Report をクリックして開くと、何本のリードがアセンブリされ、どれくらいの時間がかかり、どのようなコンティグが作成されたかがわかります。

次回は作成されたコンティグドキュメントの操作についてご紹介する予定です。

Geneious 製品概要・フリートライアルリクエストについては[こちら](#)
『Geneious Prime で猫も杓子もシーケンス解析』 過去の記事は[こちらでチェック!](#)